RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/527,672_
Source: _PG/10_
Date Processed by STIC: _3/24/05_

ENTERED

PCT

RAW SEQUENCE LISTING   DATE: 03/24/2005
PATENT APPLICATION: US/10/527,672   TIME: 11:08:16

Input Set : A:\PP019766.0003 sequence listing.TXT
Output Set: N:\CRF4\03242005\J527672.raw

```
       4 <110> APPLICANT: Rappuoli, Rino
       6 <120> TITLE OF INVENTION: GROUP B STREPTOCOCCUS VACCINE
       9 <130> FILE REFERENCE: PP019766.0003
C-->  11 <140> CURRENT APPLICATION NUMBER: US/10/527,672
C-->  11 <141> CURRENT FILING DATE: 2005-03-11
      11 <150> PRIOR APPLICATION NUMBER: PCT/US2003/029167
      12 <151> PRIOR FILING DATE: 2003-09-15
      14 <150> PRIOR APPLICATION NUMBER: US 60/410,839
      15 <151> PRIOR FILING DATE: 2002-09-13
      17 <160> NUMBER OF SEQ ID NOS: 42
      19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
      21 <210> SEQ ID NO: 1
      22 <211> LENGTH: 1662
      23 <212> TYPE: DNA
      24 <213> ORGANISM: Streptococcus agalactiae
      26 <400> SEQUENCE: 1
      27 atgaaattat cgaagaagtt attgtttтcg gctgctgttt taacaatggt ggcggggtca     60
      28 actgttgaac cagtagctca gtttgcgact ggaatgagta ttgtaagagc tgcagaagtg    120
      29 tcacaagaac gcccagcgaa aacaacagta aatatctata aattacaagc tgatagttat    180
      30 aaatcggaaa ttacttctaa tggtggtatc gagaataaag acggcgaagt aatatctaac    240
      31 tatgctaaac ttggtgacaa tgtaaaaggt ttgcaaggtg tacagtttaa acgttataaa    300
      32 gtcaagacgg atatttctgt tgatgaattg aaaaaattga caacagttga agcagcagat    360
      33 gcaaaagttg gaacgattct tgaagaaggt gtcagtctac ctcaaaaaac taatgctcaa    420
      34 ggtttggtcg tcgatgctct ggattcaaaa agtaatgtga gatacttgta tgtagaagat    480
      35 ttaaagaatt caccttcaaa cattaccaaa gcttatgctg taccgtttgt gttggaatta    540
      36 ccagttgcta actctacagg tacaggtttc ctttctgaaa ttaatattta ccctaaaaac    600
      37 gttgtaactg atgaaccaaa aacagataaa gatgttaaaa aattaggtca ggacgatgca    660
      38 ggttatacga ttggtgaaga attcaaatgg ttcttgaaat ctacaatccc tgccaattta    720
      39 ggtgactatg aaaaatttga aattactgat aaatttgcag atggcttgac ttataaatct    780
      40 gttggaaaaa tcaagattgg ttcgaaaaca ctgaatagag atgagcacta cactattgat    840
      41 gaaccaacag ttgataacca aaatacatta aaaattacgt ttaaaccaga gaaatttaaa    900
      42 gaaattgctg agctacttaa aggaatgacc cttgttaaaa atcaagatgc tcttgataaa    960
      43 gctactgcaa atacagatga tgcggcattt ttggaaattc cagttgcatc aactattaat   1020
      44 gaaaaagcag tttttaggaaa agcaattgaa aatacttttg aacttcaata tgaccatact   1080
      45 cctgataaag ctgacaatcc aaaaccatct aatcctccaa gaaaaccaga agttcatact   1140
      46 ggtgggaaac gatttgtaaa gaaagactca acagaaacac aaacactagg tggtgctgag   1200
      47 tttgatttgt tggcttctga tgggacagca gtaaatggga cagatgctct tattaaagcg   1260
      48 aatactaata aaaactatat tgctggagaa gctgttactg ggcaaccaat caaattgaaa   1320
      49 tcacatacag acggtacgtt tgagattaaa ggtttggctt atgcagttga tgcgaatgca   1380
      50 gagggtacag cagtaactta caaattaaaa gaaacaaaag caccagaagg ttatgtaatc   1440
      51 cctgataaag aaatcgagtt tacagtatca caaacatctt ataatacaaa accaactgac   1500
      52 atcacggttg atagtgctga tgcaacacct gatacaatta aaaacaacaa acgtccttca   1560
      53 atccctaata ctggtggtat tggtacggct atctttgtcg ctatcggtgc tgcggtgatg   1620
```

```
54 gcttttgctg ttaaggggat gaagcgtcgt acaaaagata ac                    1662
56 <210> SEQ ID NO: 2
57 <211> LENGTH: 554
58 <212> TYPE: PRT
59 <213> ORGANISM: Streptococcus agalactiae
61 <400> SEQUENCE: 2
62 Met Lys Leu Ser Lys Lys Leu Leu Phe Ser Ala Ala Val Leu Thr Met
63  1               5                  10                  15
64 Val Ala Gly Ser Thr Val Glu Pro Val Ala Gln Phe Ala Thr Gly Met
65              20                  25                  30
66 Ser Ile Val Arg Ala Ala Glu Val Ser Gln Glu Arg Pro Ala Lys Thr
67          35                  40                  45
68 Thr Val Asn Ile Tyr Lys Leu Gln Ala Asp Ser Tyr Lys Ser Glu Ile
69      50                  55                  60
70 Thr Ser Asn Gly Gly Ile Glu Asn Lys Asp Gly Glu Val Ile Ser Asn
71 65                  70                  75                  80
72 Tyr Ala Lys Leu Gly Asp Asn Val Lys Gly Leu Gln Gly Val Gln Phe
73              85                  90                  95
74 Lys Arg Tyr Lys Val Lys Thr Asp Ile Ser Val Asp Glu Leu Lys Lys
75          100                 105                 110
76 Leu Thr Thr Val Glu Ala Ala Asp Ala Lys Val Gly Thr Ile Leu Glu
77          115                 120                 125
78 Glu Gly Val Ser Leu Pro Gln Lys Thr Asn Ala Gln Gly Leu Val Val
79      130                 135                 140
80 Asp Ala Leu Asp Ser Lys Ser Asn Val Arg Tyr Leu Tyr Val Glu Asp
81 145                 150                 155                 160
82 Leu Lys Asn Ser Pro Ser Asn Ile Thr Lys Ala Tyr Ala Val Pro Phe
83              165                 170                 175
84 Val Leu Glu Leu Pro Val Ala Asn Ser Thr Gly Thr Gly Phe Leu Ser
85              180                 185                 190
86 Glu Ile Asn Ile Tyr Pro Lys Asn Val Val Thr Asp Glu Pro Lys Thr
87          195                 200                 205
88 Asp Lys Asp Val Lys Lys Leu Gly Gln Asp Asp Ala Gly Tyr Thr Ile
89      210                 215                 220
90 Gly Glu Glu Phe Lys Trp Phe Leu Lys Ser Thr Ile Pro Ala Asn Leu
91 225                 230                 235                 240
92 Gly Asp Tyr Glu Lys Phe Glu Ile Thr Asp Lys Phe Ala Asp Gly Leu
93              245                 250                 255
94 Thr Tyr Lys Ser Val Gly Lys Ile Lys Ile Gly Ser Lys Thr Leu Asn
95              260                 265                 270
96 Arg Asp Glu His Tyr Thr Ile Asp Glu Pro Thr Val Asp Asn Gln Asn
97          275                 280                 285
98 Thr Leu Lys Ile Thr Phe Lys Pro Glu Lys Phe Lys Glu Ile Ala Glu
99      290                 295                 300
100 Leu Leu Lys Gly Met Thr Leu Val Lys Asn Gln Asp Ala Leu Asp Lys
101 305                 310                 315                 320
102 Ala Thr Ala Asn Thr Asp Asp Ala Ala Phe Leu Glu Ile Pro Val Ala
103              325                 330                 335
104 Ser Thr Ile Asn Glu Lys Ala Val Leu Gly Lys Ala Ile Glu Asn Thr
```

Input Set : A:\PP019766.0003 sequence listing.TXT
Output Set: N:\CRF4\03242005\J527672.raw

```
105              340                 345                 350
106 Phe Glu Leu Gln Tyr Asp His Thr Pro Asp Lys Ala Asp Asn Pro Lys
107          355                 360                 365
108 Pro Ser Asn Pro Pro Arg Lys Pro Glu Val His Thr Gly Gly Lys Arg
109      370                 375                 380
110 Phe Val Lys Lys Asp Ser Thr Glu Thr Gln Thr Leu Gly Gly Ala Glu
111 385                 390                 395                 400
112 Phe Asp Leu Leu Ala Ser Asp Gly Thr Ala Val Lys Trp Thr Asp Ala
113                  405                 410                 415
114 Leu Ile Lys Ala Asn Thr Asn Lys Asn Tyr Ile Ala Gly Glu Ala Val
115              420                 425                 430
116 Thr Gly Gln Pro Ile Lys Leu Lys Ser His Thr Asp Gly Thr Phe Glu
117          435                 440                 445
118 Ile Lys Gly Leu Ala Tyr Ala Val Asp Ala Asn Ala Glu Gly Thr Ala
119      450                 455                 460
120 Val Thr Tyr Lys Leu Lys Glu Thr Lys Ala Pro Glu Gly Tyr Val Ile
121 465                 470                 475                 480
122 Pro Asp Lys Glu Ile Glu Phe Thr Val Ser Gln Thr Ser Tyr Asn Thr
123                  485                 490                 495
124 Lys Pro Thr Asp Ile Thr Val Asp Ser Ala Asp Ala Thr Pro Asp Thr
125          500                 505                 510
126 Ile Lys Asn Asn Lys Arg Pro Ser Ile Pro Asn Thr Gly Gly Ile Gly
127          515                 520                 525
128 Thr Ala Ile Phe Val Ala Ile Gly Ala Ala Val Met Ala Phe Ala Val
129      530                 535                 540
130 Lys Gly Met Lys Arg Arg Thr Lys Asp Asn
131 545                 550
```

133 <210> SEQ ID NO: 3
134 <211> LENGTH: 1629
135 <212> TYPE: DNA
136 <213> ORGANISM: Streptococcus agalactiae
138 <400> SEQUENCE: 3

```
139 atgaaaaaag gacaagtaaa tgatactaag caatcttact ctctacgtaa atataaattt      60
140 ggtttagcat cagtaatttt agggtcattc ataatggtca caagtcctgt ttttgcggat     120
141 caaactacat cggttcaagt taataatcag acaggcacta gtgtggatgc taataattct     180
142 tccaatgaga caagtgcgtc aagtgtgatt acttccaata atgatagtgt tcaagcgtct     240
143 gataaagttg taaatagtca aaatacggca acaaaggaca ttactactcc tttagtagag     300
144 acaaagccaa tggtggaaaa aacattacct gaacaaggga attatgttta tagcaaagaa     360
145 accgaggtga aaaatacacc ttcaaaatca gccccagtag ctttctatgc aaagaaaggt     420
146 gataaagttt tctatgacca agtatttaat aaagataatg tgaaatggat ttcatataag     480
147 tcttttgtg gcgtacgtcg atacgcagct attgagtcac tagatccatc aggaggttca     540
148 gagactaaag cacctactcc tgtaacaaat tcaggaagca ataatcaaga gaaaatagca     600
149 acgcaaggaa attatacatt ttcacataaa gtagaagtaa aaaatgaagc taaggtagcg     660
150 agtccaactc aatttacatt ggacaaagga gacagaattt tttacgacca aatactaact     720
151 attgaaggaa atcagtggtt atcttataaa tcattcaatg gtgttcgtcg ttttgttttg     780
152 ctaggtaaag catcttcagt agaaaaaact gaagataaag aaaaagtgtc tcctcaacca     840
153 caagcccgta ttactaaaac tggtagactg actatttcta acgaacaac tacaggtttt     900
154 gatattttaa ttacgaatat taaagatgat aacggtatcg ctgctgttaa ggtaccggtt     960
155 tggactgaac aaggagggca agatgatatt aaatggtata cagctgtaac tactggggat    1020
```

```
156 ggcaactaca aagtagctgt atcatttgct gaccataaga atgagaaggg tctttataat   1080
157 attcatttat actaccaaga agctagtggg acacttgtag gtgtaacagg aactaaagtg   1140
158 acagtagctg gaactaattc ttctcaagaa cctattgaaa atggtttagc aaagactggt   1200
159 gtttataata ttatcggaag tactgaagta aaaaatgaag ctaaaatatc aagtcagacc   1260
160 caatttactt tagaaaaagg tgacaaaata aattatgatc aagtattgac agcagatggt   1320
161 taccagtgga tttcttacaa atcttatagt ggtgttcgtc gctatattcc tgtgaaaaag   1380
162 ctaactacaa gtagtgaaaa agcgaaagat gaggcgacta aaccgactag ttatcccaac   1440
163 ttacctaaaa caggtaccta tacatttact aaaactgtag atgtgaaaag tcaacctaaa   1500
164 gtatcaagtc cagtggaatt taattttcaa aagggtgaaa aaatacatta tgatcaagtg   1560
165 ttagtagtag atggtcatca gtggatttca tacaagagtt attccggtat tcgtcgctat   1620
166 attgaaatt                                                           1629
168 <210> SEQ ID NO: 4
169 <211> LENGTH: 543
170 <212> TYPE: PRT
171 <213> ORGANISM: Streptococcus agalactiae
173 <400> SEQUENCE: 4
174 Met Lys Lys Gly Gln Val Asn Asp Thr Lys Gln Ser Tyr Ser Leu Arg
175  1               5                  10                  15
176 Lys Tyr Lys Phe Gly Leu Ala Ser Val Ile Leu Gly Ser Phe Ile Met
177         20                  25                  30
178 Val Thr Ser Pro Val Phe Ala Asp Gln Thr Thr Ser Val Gln Val Asn
179         35                  40                  45
180 Asn Gln Thr Gly Thr Ser Val Asp Ala Asn Asn Ser Ser Asn Glu Thr
181     50                  55                  60
182 Ser Ala Ser Ser Val Ile Thr Ser Asn Asn Asp Ser Val Gln Ala Ser
183 65                  70                  75                  80
184 Asp Lys Val Val Asn Ser Gln Asn Thr Ala Thr Lys Asp Ile Thr Thr
185                 85                  90                  95
186 Pro Leu Val Glu Thr Lys Pro Met Val Glu Lys Thr Leu Pro Glu Gln
187                 100                 105                 110
188 Gly Asn Tyr Val Tyr Ser Lys Glu Thr Glu Val Lys Asn Thr Pro Ser
189         115                 120                 125
190 Lys Ser Ala Pro Val Ala Phe Tyr Ala Lys Lys Gly Asp Lys Val Phe
191     130                 135                 140
192 Tyr Asp Gln Val Phe Asn Lys Asp Asn Val Lys Trp Ile Ser Tyr Lys
193 145                 150                 155                 160
194 Ser Phe Cys Gly Val Arg Arg Tyr Ala Ala Ile Glu Ser Leu Asp Pro
195                 165                 170                 175
196 Ser Gly Gly Ser Glu Thr Lys Ala Pro Thr Pro Val Thr Asn Ser Gly
197                 180                 185                 190
198 Ser Asn Asn Gln Glu Lys Ile Ala Thr Gln Gly Asn Tyr Thr Phe Ser
199         195                 200                 205
200 His Lys Val Glu Val Lys Asn Glu Ala Lys Val Ala Ser Pro Thr Gln
201     210                 215                 220
202 Phe Thr Leu Asp Lys Gly Asp Arg Ile Phe Tyr Asp Gln Ile Leu Thr
203 225                 230                 235                 240
204 Ile Glu Gly Asn Gln Trp Leu Ser Tyr Lys Ser Phe Asn Gly Val Arg
205                 245                 250                 255
206 Arg Phe Val Leu Leu Gly Lys Ala Ser Ser Val Glu Lys Thr Glu Asp
```

RAW SEQUENCE LISTING   DATE: 03/24/2005
PATENT APPLICATION: US/10/527,672   TIME: 11:08:16

Input Set : A:\PP019766.0003 sequence listing.TXT
Output Set: N:\CRF4\03242005\J527672.raw

```
207              260              265              270
208 Lys Glu Lys Val Ser Pro Gln Pro Gln Ala Arg Ile Thr Lys Thr Gly
209         275              280              285
210 Arg Leu Thr Ile Ser Asn Glu Thr Thr Thr Gly Phe Asp Ile Leu Ile
211     290              295              300·
212 Thr Asn Ile Lys Asp Asp Asn Gly Ile Ala Ala Val Lys Val Pro Val
213 305              310              315              320
214 Trp Thr Glu Gln Gly Gly Gln Asp Asp Ile Lys Trp Tyr Thr Ala Val
215              325              330              335
216 Thr Thr Gly Asp Gly Asn Tyr Lys Val Ala Val Ser Phe Ala Asp His
217         340              345              350
218 Lys Asn Glu Lys Gly Leu Tyr Asn Ile His Leu Tyr Tyr Gln Glu Ala
219         355              360              365
220 Ser Gly Thr Leu Val Gly Val Thr Gly Thr Lys Val Thr Val Ala Gly
221     370              375              380
222 Thr Asn Ser Ser Gln Glu Pro Ile Glu Asn Gly Leu Ala Lys Thr Gly
223 385              390              395              400
224 Val Tyr Asn Ile Ile Gly Ser Thr Glu Val Lys Asn Glu Ala Lys Ile
225              405              410              415
226 Ser Ser Gln Thr Gln Phe Thr Leu Glu Lys Gly Asp Lys Ile Asn Tyr
227         420              425              430
228 Asp Gln Val Leu Thr Ala Asp Gly Tyr Gln Trp Ile Ser Tyr Lys Ser
229         435              440              445
230 Tyr Ser Gly Val Arg Arg Tyr Ile Pro Val Lys Lys Leu Thr Thr Ser
231     450              455              460
232 Ser Glu Lys Ala Lys Asp Glu Ala Thr Lys Pro Thr Ser Tyr Pro Asn
233 465              470              475              480
234 Leu Pro Lys Thr Gly Thr Tyr Thr Phe Thr Lys Thr Val Asp Val Lys
235              485              490              495
236 Ser Gln Pro Lys Val Ser Ser Pro Val Glu Phe Asn Phe Gln Lys Gly
237         500              505              510
238 Glu Lys Ile His Tyr Asp Gln Val Leu Val Val Asp Gly His Gln Trp
239         515              520              525
240 Ile Ser Tyr Lys Ser Tyr Ser Gly Ile Arg Arg Tyr Ile Glu Ile
241     530              535              540
243 <210> SEQ ID NO: 5
244 <211> LENGTH: 2670
245 <212> TYPE: DNA
246 <213> ORGANISM: Streptococcus agalactiae
248 <400> SEQUENCE: 5
249 atgaaaaaga gacaaaaaat atggagaggg ttatcagtta ctttactaat cctgtcccaa      60
250 attccatttg gtatattggt acaaggtgaa acccaagata ccaatcaagc acttggaaaa     120
251 gtaattgtta aaaaaacggg agacaatgct acaccattag caaagcgac ttttgtgtta     180
252 aaaaatgaca atgataagtc agaaacaagt cacgaaacgg tagagggttc tggagaagca     240
253 acctttgaaa acataaaacc tggagactac acattaagag aagaaacagc accaattggt     300
254 tataaaaaaa ctgataaaac ctggaaagtt aaagttgcag ataacggagc aacaataatc     360
255 gagggtatgg atgcagataa agcagagaaa cgaaaagaag ttttgaatgc ccaatatcca     420
256 aaatcagcta tttatgagga tacaaaagaa aattacccat tagttaatgt agagggttcc     480
257 aaagttggtg aacaatacaa agcattgaat ccaataaatg gaaaagatgg tcgaagagag     540
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 03/24/2005
PATENT APPLICATION: US/10/527,672   TIME: 11:08:17

Input Set : A:\PP019766.0003 sequence listing.TXT
Output Set: N:\CRF4\03242005\J527672.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; Xaa Pos. 240
Seq#:40; Xaa Pos. 281

VERIFICATION SUMMARY　　　　　　　　　DATE: 03/24/2005
PATENT APPLICATION: US/10/527,672　　　TIME: 11:08:17

Input Set : **A:\PP019766.0003 sequence listing.TXT**
Output Set: **N:\CRF4\03242005\J527672.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1462 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:224
L:2167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:272